

# Systematic DFT Frames: Principle, Eigenvalues Structure, and Applications

Mojtaba Vaezi and Fabrice Labeau

## Abstract

Motivated by a host of recent applications requiring some amount of redundancy, *frames* are becoming a standard tool in the signal processing toolbox. In this paper, we study a specific class of frames, known as discrete Fourier transform (DFT) codes, and introduce the notion of *systematic* frames for this class. This is encouraged by a new application of frames, namely, distributed source coding that uses DFT codes for compression. Studying their extreme eigenvalues, we show that, unlike DFT frames, systematic DFT frames are not necessarily *tight*. Then, we come up with conditions for which these frames can be tight. In either case, the best and worst systematic frames are established in the minimum mean-squared reconstruction error sense. Eigenvalues of DFT frames and their subframes play a pivotal role in this work. Particularly, we derive some bounds on the extreme eigenvalues DFT subframes which are used to prove most of the results; these bounds are valuable independently.

## Index Terms

BCH-DFT codes, frames, systematic frames, parity, eigenvalue, optimal reconstruction, quantization, erasures, distributed source coding, Vandermonde matrix.

## I. INTRODUCTION

Frames, “redundant” set of vectors used for signal representation, are increasingly found in signal processing applications. Frames are more general than bases as frames are complete but not necessarily

This work was supported by Hydro-Québec, the Natural Sciences and Engineering Research Council of Canada and McGill University in the framework of the NSERC/Hydro-Québec/McGill Industrial Research Chair in Interactive Information Infrastructure for the Power Grid. Part of the material in this paper was presented at the International Symposium on Information Theory, Boston, MA, July 2012 [1].

The authors are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: mojtava.vaezi@mail.mcgill.ca; fabrice.labeau@mcgill.ca).

linearly independent. A *basis*, on the contrary, is a set of vectors used to “uniquely” represent a vector as a linear combination of basis elements. What would be the benefit of representing a signal with overcomplete set of vectors? Frames are generally motivated by applications requiring some level of redundancy. They offer flexibility in design and have a variety of applications. Frames show resilience to additive noise (including quantization error), robustness to erasure (loss), and numerical stability of reconstruction. With increasing applications, frames are becoming more prevalent in signal processing.

In this paper, we study a specific class of frames known as discrete Fourier transform (DFT) codes. By using these codes, the ideas of coding theory are described in the signal processing setting. We consider the Bose-Chaudhuri-Hocquenghem (BCH) codes, an important class of multiple-error-correcting codes, in the DFT domain [2]–[4]. BCH-DFT codes are *cyclic codes* in the complex (or real) domain, similar to BCH codes in the binary error correction setting. Therefore, their codewords have certain successive spectral components equal to zero. This property is then exploited for error detection and correction in the complex (real) field [2]–[9].

From Frame theory perspective, DFT codes are harmonic tight frames. In the absence of erasure, tight frames minimize the mean-squared error (MSE) between the transmitted and received signals [10]–[12]. The MSE is the ultimate measure of performance in many digital communication systems where quantized analog signal is transmitted. Frames are naturally robust to transmission loss since they provide an overcomplete expansion of signal [10]–[14].

DFT frames have recently been applied in the context of distributed source coding (DSC) [15]. More precisely, BCH-DFT codes are used for compression of analog signals with side information available at the decoder. In DSC context, side information is viewed as corrupted version of signal, and compression is achieved by sending only redundant information, in the form of *parity* or *syndrome*, with respect to a channel code [16]. Unlike in DSC that uses binary channel codes for compression, in the new framework (DSC based on BCH-DFT codes) compression is performed before quantization. As a result, DFT frames, which are primarily used for compression, can decrease quantization error at the same time. This results in a better reconstruction, in the MSE sense, particularly when the sources are highly correlated.

Motivated by its application in parity-based DSC that uses DFT codes [15], we introduce the notion of *systematic frames*, in this work. For an  $(n, k)$  frame, a systematic frame is defined to be a frame that includes the identity matrix of size  $k$  as a subframe. Since *tight* frames minimize reconstruction error [10]–[13], we explore *systematic tight DFT frames*. Although it is straightforward to construct systematic DFT frames, we prove that systematic “tight” DFT frames exist only for specific DFT codes. More precisely, we show that a systematic frame is tight if and only if data (systematic) samples are circularly

equally spaced, in the codewords generated by that frame. When such a frame does not exist, we will be looking for systematic DFT frames with the “best” performance, from the minimum mean-squared reconstruction error sense. We also demonstrate which systematic frames are the “worst” in this sense. In addition, we show that circular shift and reversal of the vectors in a DFT frame does not change the eigenvalues of the frame operator. We use these properties to categorize different systematic frames of an  $(n, k)$  DFT frame based on their performance.

Another main contribution of this paper is to find bounds on the *extreme eigenvalues* of  $V^H V$ , where  $V$  is a square or non-square subframe of a DFT frame. The properties of the eigenvalues of such frames are central to establish many of the result in this paper. These bounds are used to determine the conditions required for a systematic frame so as to be tight. Besides, eigenvalues are crucial in establishing the best and worst systematic frames.

The paper is organized as follows. In Section II, we present the basic definitions and a few fundamental lemmas that will be used later in the paper. In Section III, we introduce DFT frames and set the ground to study the extreme eigenvalues of their subframes. Section IV motivates the work in this paper by introducing systematic DFT frames and their application. Some result on the the extreme eigenvalues of DFT frames and their subframes are presented in Section V. Sections VI and VII is devoted to the evaluation of reconstruction error and classification of systematic frames based on that. We conclude in Section VIII.

For notation, we use boldface lower-case letters for vectors, boldface upper-case letters for matrices,  $(\cdot)^T$  for transpose,  $(\cdot)^H$  for conjugate transpose,  $(\cdot)^\dagger$  for pseudoinverse,  $(\cdot)^*$  for conjugate,  $\text{tr}(\cdot)$  for the trace,  $\mathbb{E}(\cdot)$  for the mathematical expectation, and  $\|\cdot\|$  for the Euclidean norm. The dimensions of square and rectangular matrices are indicated, respectively, by one and two subscripts when required.

## II. DEFINITIONS AND PRELIMINARIES

In this section, we introduce the definitions and some basic results which are frequently used in the paper.

**Definition 1.** A spanning family of  $n$  vectors  $F = \{\mathbf{f}_i\}_{i=1}^n$  in a  $k$ -dimensional complex vector space  $\mathbb{C}^k$  is called a *frame* if there exist  $0 < a \leq b$  such that for any  $\mathbf{x} \in \mathbb{C}^k$

$$a\|\mathbf{x}\|^2 \leq \sum_{i=1}^n |\langle \mathbf{x}, \mathbf{f}_i \rangle|^2 \leq b\|\mathbf{x}\|^2, \quad (1)$$

where  $\langle \mathbf{x}, \mathbf{f}_i \rangle$  denotes the inner product of  $\mathbf{x}$  and  $\mathbf{f}_i$  and gives the  $i$ th coefficient for the frame expansion of  $\mathbf{x}$  [12]–[14].  $a$  and  $b$  are called *frame bounds*; they, respectively, ensure that the vectors span the space,

and the basis expansion converges. A frame is *tight* if  $a = b$ . *Uniform* or *equal-norm* frames are frames with same norm for all elements, i.e.,  $\|\mathbf{f}_i\| = \|\mathbf{f}_j\|$ , for  $i, j = 1, \dots, n$ .

**Definition 2.** An  $n \times n$  Vandermonde matrix with unit complex entries is defined by

$$W \triangleq \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \dots & e^{j\theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(n-1)\theta_1} & e^{j(n-1)\theta_2} & \dots & e^{j(n-1)\theta_n} \end{pmatrix}, \quad (2)$$

in which  $\theta_p \in [0, 2\pi)$  and  $\theta_p \neq \theta_q$  for  $p \neq q$ ,  $1 \leq p, q \leq n$ . If  $\theta_p = \frac{2\pi}{n}(p-1)$ ,  $W$  becomes the well-known IDFT matrix [17]. For this Vandermonde matrix we can write [18], [19]

$$\det(WW^H) = |\det(W)|^2 = \frac{1}{n^n} \prod_{1 \leq p < q \leq n} |e^{i\theta_p} - e^{i\theta_q}|^2. \quad (3)$$

Central to this work is the properties of the *eigenvalues* of  $V^H V$  or  $VV^H$ , in which  $V$  is a submatrix of a DFT matrix.<sup>1</sup> Hence, we recall some bounds on the eigenvalues of Hermitian matrices which are used in this paper. Let  $A$  be a Hermitian  $k \times k$  matrix with real eigenvalues  $\{\lambda_1(A), \dots, \lambda_k(A)\}$  which are collectively called the *spectrum* of  $A$ , and assume  $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_k(A)$ . Schur-Horn inequalities show to what extent the eigenvalues of a Hermitian matrix constraint its diagonal entries.

**Proposition 1.** *Schur-Horn inequalities* [20]

Let  $A$  be a Hermitian  $k \times k$  matrix with real eigenvalues  $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_k(A)$ . Then, for any  $1 \leq i_1 < i_2 < \dots < i_l \leq k$ ,

$$\begin{aligned} \lambda_{k-l+1}(A) + \dots + \lambda_k(A) &\leq a_{i_1 i_1} + \dots + a_{i_l i_l} \\ &\leq \lambda_1(A) + \dots + \lambda_l(A), \end{aligned} \quad (4)$$

where  $a_{11}, \dots, a_{kk}$  are the diagonal entries of  $A$ . Particularly, for  $l = 1$  and  $l = k$  we obtain

$$\lambda_k(A) \leq a_{11} \leq \lambda_1(A), \quad (5)$$

$$\sum_{i=1}^k \lambda_i(A) = \sum_{i=1}^k a_{ii}. \quad (6)$$

<sup>1</sup>Note that eigenvalues of  $V^H V$  and  $VV^H$  are equal for a square  $V$ ; also,  $V^H V$  and  $VV^H$  have the same nonzero eigenvalues for a non-square  $V$ .

Another basic question in linear algebra asks the degree to which the eigenvalues of two Hermitian matrices constrain the eigenvalues of their sum. Weyl's theorem gives an answer to this question in the following set of inequalities.

**Proposition 2.** *Weyl inequalities [20]*

Let  $A$  and  $B$  be two Hermitian  $k \times k$  matrices with spectrums  $\{\lambda_1(A), \dots, \lambda_k(A)\}$  and  $\{\lambda_1(B), \dots, \lambda_k(B)\}$ , respectively. Then, for  $i, j \leq k$ , we have

$$\lambda_i(A + B) \leq \lambda_j(A) + \lambda_{i-j+1}(B) \quad \text{for } j \leq i, \quad (7)$$

$$\lambda_i(A + B) \geq \lambda_j(A) + \lambda_{k+i-j}(B) \quad \text{for } j \geq i. \quad (8)$$

**Corollary 1.** If  $A + B = \gamma I_k$ ,  $\gamma > 0$ , where  $A$  and  $B$  are Hermitian matrices, then  $\lambda_j(A) + \lambda_{k-j+1}(B) = \gamma$ .

*Proof:* It suffice to set  $i = k$  and  $i = 1$  respectively in (7) and (8), and use  $\lambda_k(A + B) = \lambda_1(A + B) = \gamma$  which is obtained from  $A + B = \gamma I_k$ . ■

**Lemma 1.** Let  $A$  and  $B$  be two Hermitian  $k \times k$  matrices and suppose that, for every  $1 \leq i, j \leq k$ ,  $A_{i,j} = e^{j\theta_i} B_{i,j}$ ; then  $A^H A$  and  $B^H B$  have the same spectrum.

*Proof:* The proof is immediate using Lemma 3 [19] since  $(A^H A)_{i,j} = \frac{e^{j\theta_i}}{e^{j\theta_i}} (B^H B)_{i,j}$ ; i.e.,  $A^H A = B^H B$ . ■

### III. DFT FRAMES

#### A. BCH-DFT Codes

DFT codes [3], are linear block codes over the complex field whose parity-check matrix  $H$  is defined based on the DFT matrix. A Bose-Chaudhuri-Hocquenghem (BCH) DFT code is a DFT code that insert  $d - 1$  cyclically adjacent zeros in the frequency-domain function (Fourier transform) of any codeword where  $d$  is the designed distance of that code [2]. Real BCH-DFT codes, a subset of complex BCH-DFT codes, benefit from a generator matrix with real entries. The generator matrix of an  $(n, k)$  real BCH-DFT code<sup>2</sup> is typically defined by [6], [8], [11], [15]

$$G = \sqrt{\frac{n}{k}} W_n^H \Sigma W_k, \quad (9)$$

<sup>2</sup>Real BCH-DFT codes do not exist [3] when  $n$  and  $k$  are simultaneously even.

in which  $W_l$  represents the DFT matrix of size  $l$ , and  $\Sigma$  is defined as

$$\Sigma_{n \times k} = \begin{pmatrix} I_\alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_\beta \end{pmatrix}, \quad (10)$$

where  $\alpha = \lceil n/2 \rceil - \lfloor (n-k)/2 \rfloor$ ,  $\beta = k - \alpha$ , and the sizes of zero blocks are such that  $\Sigma$  is an  $n \times k$  matrix [8]. One can check that  $\Sigma^H \Sigma = I_k$ , and  $\Sigma \Sigma^H$  is an  $n \times n$  matrix given by

$$\Sigma \Sigma^H = \begin{pmatrix} I_\alpha & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_\beta \end{pmatrix}. \quad (11)$$

The above structure guarantees a real BCH code. Note that, having  $n-k$  consecutive zero rows,  $\Sigma$  inserts  $n-k$  consecutive zeros to each codeword in the frequency domain, which is required to make a BCH code [2].

*Remark 1.* Removing  $W_k$  from (9) we end up with a complex  $G$ , representing a *complex BCH-DFT code*. In such a code,  $\alpha$  and  $\beta$  can be any nonnegative integers such that  $\alpha + \beta = k$ .

The parity-check matrix  $H$ , both in real and complex codes, consist of the  $n-k$  columns of  $W_n^H$  corresponding to the zero rows of  $\Sigma$ ; thus,  $HG = 0$ .

### B. Connection to Frame Theory

The generator matrix  $G$  in (9) can be viewed as an *analysis frame operator*. In this view, a real BCH-DFT code is a rotation of the well-known *harmonic frames* [13], [14], and a complex BCH-DFT code is basically a harmonic frame. The latter can be understood by removing  $W_k$  from (9) which results in a complex BCH-DFT code, on the one hand, and the analysis frame operator of a harmonic frame, on the other hand. The former is then evident as  $W_k$  is a rotation matrix. Further, it is easy to see that the *frame operator*  $G^H G$  and *Gramian*  $GG^H$  are equal to

$$G^H G = \frac{n}{k} I_k, \quad (12)$$

$$GG^H = \frac{n}{k} W_n^H \Sigma \Sigma^H W_n. \quad (13)$$

The following lemma presents some properties of the frame operator and relevant matrices which are crucial for our results in this paper.

**Lemma 2.** *Let  $G_{p \times k}$  be a matrix consisting of  $p$  arbitrary rows of  $G$  defined by (9). Then, the following statements hold:*

- i.  $GG^H$  is a Toeplitz and circulant matrix
- ii.  $G_{p \times k} G_{p \times k}^H$ ,  $1 < p < n$  is a Toeplitz matrix
- iii. All principal diagonal entries of  $G_{p \times k} G_{p \times k}^H$ ,  $1 \leq p \leq n$  are equal to 1.

*Proof:* Let  $a_{r,s}$  be the  $(r, s)$  entry of the matrix  $GG^H$  then it can readily be shown that

$$a_{r,s} = \frac{1}{k} \sum_{m=0}^{\alpha-1} e^{jm(\theta_r - \theta_s)} + \frac{1}{k} \sum_{m=n-\beta}^{n-1} e^{jm(\theta_r - \theta_s)}, \quad (14)$$

in which  $\theta_x = \frac{2\pi}{n}(x-1)$ . From this equation, it is clear that  $a_{r,s} = a_{r+i, s+i}$ ; that is, the elements of each diagonal are equal, which means that  $GG^H$  is a Toeplitz matrix. In addition, we can check that  $a_{r,n} = a_{r+1,1}$ , i.e., the last entry in each row is equal to the first entry of the next row. This proves that the Toeplitz matrix  $GG^H$  is circulant as well [21]. Also, a quick look at (14) reveals that the elements of the principal diagonal ( $r = s$ ) are equal to 1. Similarly, one can see that for any  $1 < p < n$ , the square matrix  $G_{p \times k} G_{p \times k}^H$  is also a Toeplitz matrix; it is not necessarily circulant, however. ■

Considering Remark 1, one can check that (14) is also valid for complex BCH-DFT codes. Note that,  $\alpha$  and  $\beta$  are less constrained for these codes, as mentioned in Remark 1.

*Remark 2.* Lemma 2 also holds for complex BCH-DFT codes.

Further, in a DFT code, in general, the  $n - k$  zero rows of  $\Sigma$  are not required to be successive if they are not designed for error correction. That is any matrix that can be rearranged as  $[I_k \mid \mathbf{0}_{k \times n-k}]^T$  may represent  $\Sigma$ . Then,  $\Sigma \Sigma^H$  is not necessarily in the form given in (11); it can be any square matrix of size  $n$  with  $k$  nonzero elements equal to 1, arbitrary located on the main diagonal. Then, again Lemma 2 holds because  $a_{rs} = \frac{1}{k} \sum_{i=0}^{k-1} e^{jm_i(\theta_r - \theta_s)}$  and  $m_i \in \{1, \dots, n\}$ .

*Remark 3.* Lemma 2 holds for all DFT codes.

As a result, all properties that we prove in the remainder of this paper are valid for “any” DFT code, even though they are formally proved for real BCH-DFT codes, which we simply refer to as “DFT codes” or “DFT frames” hereafter.

## IV. SYSTEMATIC DFT FRAMES

### A. Motivation

In the context of channel coding, there is a special interest in *systematic codes* [2] since the input data is embedded in the encoded output which simplifies the encoding and decoding algorithms. For example,

in systematic convolutional codes data can be read directly if no errors are made, or in case errors occur in erasures and affect only parity bits.

Systematic codes are also used in parity-based distributed source coding (DSC) techniques, e.g., DSC that uses turbo codes for compression [22]–[24]. DSC addresses the problem of compressing correlated sources by separate encoding and joint decoding and has found application in sensor networks and video compression [16]. In DSC, compression is usually realized through the use of *binary channel codes*.

Recently, the authors has introduced a new framework that uses *real-number codes* for DSC. Specifically, by using BCH-DFT codes it has been shown that this framework can result a better compression compared to the conventional one [15]. Although parity-based and syndrome-based approaches are equal in a DSC that uses binary channel codes for compression, the parity-based approach is more worthwhile in the DSC that uses DFT codes, as it is more “efficient” than the syndrome-based approach [15]. This is due to the fact that syndrome is a complex vector, even in a real DFT code, whereas parity is real. The parity-based approach requires a systematic generator matrix for DFT codes and is the main motivation of this work.

### B. Construction

A systematic generator matrix for a real BCH-DFT code can be obtained by [15]

$$G_{\text{sys}} = GG_k^{-1}, \quad (15)$$

in which  $G_k$  is a submatrix (subframe) of  $G$  including  $k$  arbitrary rows of  $G$ . Note that  $G_k$  is invertible since it can be represented as

$$G_k = \sqrt{\frac{n}{k}} W_{k \times n}^H \Sigma W_k = V_k^H W_k, \quad (16)$$

in which  $V_k^H \triangleq \sqrt{\frac{n}{k}} W_{k \times n}^H \Sigma$  and  $W_k$  are invertible as they are Vandermonde and DFT matrices, respectively. This indicates that any  $k$  rows of a DFT frame make a frame and proves that systematic DFT frames exist for any DFT frame. It also suggests that, for each DFT frame, there are many (but, a finite number of) systematic frames since the rows of  $G_k$  can be arbitrarily chosen from those of  $G$ . This will be discussed in detail later in Section VII-C. The codewords generated by these systematic frames differ in the “position” of systematic samples (i.e., input data).

The construction in (15) implies that the parity samples are not restricted to form a consecutive block in the associated codewords. Such a degree of freedom is useful in the sense that one can find the most suitable systematic frames for specific applications (e.g., the one with the smallest reconstruction error.)



### C. Optimality Condition

From rate-distortion theory, it is well known that the rate required to transmit a source, with a given distortion, increases as the variance of the source becomes larger [25]. Particularly, for Gaussian sources this relation is logarithmic with variance, under the mean-squared error (MSE) distortion measure. In DSC that uses real-number codes [15], since coding is performed before quantization, the variance of transmitted sequence depends on the behavior of the encoding matrix. In syndrome approach,  $\mathbf{s} = H\mathbf{x}$  [15] and it can be checked that  $\sigma_s = \sigma_x$ , that is, the variance is preserved.<sup>3</sup> However, as we show shortly, this is not valid in parity approach and the variance of parity samples depends on the behavior of encoding matrix  $G_{\text{sys}}$ . In view of rate-distortion theory, it makes a lot of sense to keep this variance as small as possible. Not surprisingly, we will show that using a tight frame (tight  $G_{\text{sys}}$ ) for encoding is optimal.

Let  $\mathbf{x}$  be the message vector, a column vector whose elements are i.i.d. random variables with variance  $\sigma_x^2$ , and let  $\mathbf{y} = G_{\text{sys}}\mathbf{x}$  represent the codeword generated using the systematic frame. The variance of  $\mathbf{y}$  is then given by

$$\begin{aligned}\sigma_y^2 &= \frac{1}{n} \mathbb{E}\{\mathbf{y}^H \mathbf{y}\} = \frac{1}{n} \mathbb{E}\{\mathbf{x}^H G_{\text{sys}}^H G_{\text{sys}} \mathbf{x}\} \\ &= \frac{1}{n} \sigma_x^2 \text{tr}(G_{\text{sys}}^H G_{\text{sys}}),\end{aligned}\tag{17}$$

and

$$\begin{aligned}\text{tr}(G_{\text{sys}}^H G_{\text{sys}}) &= \text{tr}\left(G_k^{-1H} G^H G G_k^{-1}\right) \\ &= \frac{n}{k} \text{tr}\left((G_k G_k^H)^{-1}\right) \\ &= \frac{n}{k} \text{tr}\left((V_k^H V_k)^{-1}\right) \\ &= \frac{n}{k} \sum_{i=1}^k \frac{1}{\lambda_i},\end{aligned}\tag{18}$$

in which  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$  are the eigenvalues of  $G_k G_k^H$  (or  $V_k^H V_k$  equivalently).

This shows that the variance of codewords, generated by a systematic frame, depends on the submatrix  $G_k$  which is used to create  $G_{\text{sys}}$ .  $G_k$ , in turn, is fully known once the position of systematic (data) samples is fixed in the codeword. In other words, the “position” of systematic samples determines the variance of

<sup>3</sup>In general, any unitary matrix  $U$  preserves norms, i.e., for any complex vector  $\mathbf{x}$ ,  $\|U\mathbf{x}\| = \|\mathbf{x}\|$ . Note that  $H$  is not unitary because it is not a square matrix; however, its rows are selected from a unitary matrix and are orthonormal. This lead to  $HH^H = I_{n-k}$ , and  $\text{tr}(H^H H) = n - k$ .

codewords generated by a systematic DFT frame. To minimize the effective range of transmitted signal, we need to do the following optimization problem, in light of (17) and (18),

$$\begin{aligned} & \underset{\lambda_i}{\text{minimize}} && \sum_{i=1}^k \frac{1}{\lambda_i} \\ & \text{s.t.} && \sum_{i=1}^k \lambda_i = k, \lambda_i > 0, \end{aligned} \tag{19}$$

where, the constraint  $\sum_{i=1}^k \lambda_i = k$  is achieved in consideration of Lemma 2 and (6).

By using the Lagrangian method, we can show that the optimal eigenvalues are  $\lambda_i = 1$ ; this implies a tight frame [10]. In the sequel, we analyze the eigenvalues of  $G_{p \times k} G_{p \times k}^H$ ,  $p \leq n$ , that helps us characterize tight systematic frames, so as to minimize the variance of transmitted codewords.

## V. MAIN RESULTS ON THE EXTREME EIGENVALUES

In this section we investigate some bounds on the eigenvalues of  $G_{p \times k} G_{p \times k}^H$  where  $G$  is defined in (9). These bounds play an important role in the performance evaluation of the systematic DFT frames. We also determine the exact values of some eigenvalues in certain cases.

**Theorem 1.** *Let  $G_{p \times k}$ ,  $1 \leq p \leq n$  be any  $p \times k$  submatrix of  $G$ . Then, the smallest eigenvalue of  $G_{p \times k} G_{p \times k}^H$  is no more than one, and the largest eigenvalue of  $G_{p \times k} G_{p \times k}^H$  is at least one.*

*Proof:* From Lemma 2, we know that all principal diagonal entries of  $G_{p \times k} G_{p \times k}^H$  are unity. As a result, using the Schur-Horn inequality in (5), we obtain  $\lambda_{\min}(G_{p \times k} G_{p \times k}^H) \leq 1 \leq \lambda_{\max}(G_{p \times k} G_{p \times k}^H)$ . This proves the claim.  $\blacksquare$

Note that  $\lambda_1(G_{p \times k} G_{p \times k}^H) = \lambda_1(G_{p \times k}^H G_{p \times k})$  for any  $G_{p \times k}$ . Nevertheless, this is not correct for  $\lambda_{\min}$  in general. A tighter bound on  $\lambda_1$  can be achieved when  $G_{p \times k}$  is a tall<sup>4</sup> matrix.

**Theorem 2.** *Given a tall (short)  $G_{p \times k}$ , the largest (smallest) eigenvalue of  $G_{p \times k}^H G_{p \times k}$  is lower (upper) bounded by  $p/k$ .*

*Proof:* Let  $p > k$ . Since all diagonal entries of  $G_{p \times k} G_{p \times k}^H$  are unity, from (6) we have  $\sum_{i=1}^p \lambda_i(G_{p \times k} G_{p \times k}^H) = p$ . On the other hand, since the nonzero eigenvalues of  $G_{p \times k} G_{p \times k}^H$  and  $G_{p \times k}^H G_{p \times k}$  are equal,  $G_{p \times k} G_{p \times k}^H$

<sup>4</sup>An  $m \times n$  matrix  $A$  is called to be tall if  $m > n$ . Similarly, if  $m < n$ , then  $A$  is a short matrix.

has  $k$  nonzero eigenvalues and we get

$$\begin{aligned}
 p &= \sum_{i=1}^p \lambda_i(G_{p \times k} G_{p \times k}^H) \\
 &= \sum_{i=1}^k \lambda_i(G_{p \times k}^H G_{p \times k}) \\
 &\leq k \lambda_1(G_{p \times k}^H G_{p \times k}).
 \end{aligned} \tag{20}$$

Thus, for any tall  $G_{p \times k}$ ,  $\lambda_1(G_{p \times k}^H G_{p \times k}) = \lambda_1(G_{p \times k} G_{p \times k}^H) \geq \frac{p}{k} > 1$ . Following a similar line of proof, for a short submatrix ( $p < k$ ) we obtain  $\lambda_{\min}(G_{p \times k}^H G_{p \times k}) \leq \frac{p}{k} < 1$ .

Obviously the same bounds are valid for the extreme eigenvalues of  $G_{p \times k} G_{p \times k}^H$ . What is more, since  $p/k$  is the average value of eigenvalues, considering that  $\lambda_{\min}(G_{p \times k} G_{p \times k}^H) = 0$  for  $p > k$ , and  $\lambda_{\min}(G_{p \times k}^H G_{p \times k}) = 0$  for  $p < k$ , from (20) we conclude that corresponding bounds on the largest eigenvalues are strict. ■

It is worth noting that in (20) the equality is achieved when  $p = n$ ; it can also be achieved for “specific” submatrices only in the case of integer oversampling, i.e., when  $n = Mk$ , as we discuss later in this paper.

We use the above results to find better bounds for the extreme eigenvalues of  $G_k G_k^H$  in the following theorem.

**Theorem 3.** *For any  $G_k$ , a square submatrix of  $G$  in (9) in which  $n \neq Mk$ , the smallest (largest) eigenvalue of  $G_k G_k^H$  is strictly upper (lower) bounded by 1.*

*Proof:* See Appendix IX-A. ■

Theorem 3 implies that for  $n \neq Mk$  we cannot have “tight” systematic frames. Because, for a frame with frame operator  $F^H F$ , the tightest possible frame bounds are, respectively,  $a = \lambda_{\min}(F^H F)$  and  $b = \lambda_{\max}(F^H F)$  [26]. In other words, for a tight frame  $\lambda_{\min}(F^H F) = \lambda_{\max}(F^H F)$ ; i.e., the eigenvalues of  $F^H F$  are equal [10].

**Corollary 2.** *Tight systematic DFT frames can exist only if  $n = Mk$ , where  $M$  is a positive integer.*

Note that systematic DFT frames are not necessarily tight for  $n = Mk$ . In Section VII, we prove that tight systematic DFT frames exist for  $n = Mk$  and show how to construct such frames.

In the remainder of this section, we shall find exact values, rather than bounds, for some of the eigenvalues of  $G_k^H G_k$  when  $k < n \leq 2k$ . This range of  $n$  is specifically important in distributed source coding.

**Theorem 4.** For any  $G_k$ , a square submatrix of  $G$  in (9), where  $k < n < 2k$ , the  $2k - n$  largest eigenvalues of  $G_k G_k^H$  are equal to  $n/k$ .

*Proof:* From Corollary 1 we know that if two Hermitian matrices sum up to a scaled identity matrix, their eigenvalues add up to be fixed. Thus, if  $A$  and  $B$  have the same spectrum we obtain

$$\lambda_j(A) + \lambda_{k-j+1}(A) = \gamma. \quad (21)$$

Now, let  $G$  be partitioned as  $G = \begin{bmatrix} G_k \\ \bar{G}_{p \times k} \end{bmatrix}$  where  $p = n - k$ . Let  $A = G_k^H G_k$  and  $B = \bar{G}_{p \times k}^H \bar{G}_{p \times k}$ , then  $A + B = G^H G = \frac{n}{k} I_k$ . Clearly, Corollary 1 holds with  $\gamma = \frac{n}{k}$ . Also, note that when  $p < k$  then  $\bar{G}_{p \times k}^H \bar{G}_{p \times k}$  has only  $p$  nonzero eigenvalues. Therefore, in such a case,  $k - p$  largest eigenvalues of  $G_k^H G_k$  are equal to  $n/k$ . ■

Another interesting case arises when  $n = 2k$ . Numerical results shows that under this condition,  $A$  and  $B$  have the same set of eigenvalues. We prove this when  $G_k$  either includes successive or every other rows of  $G$ . In such cases, one can verify that  $(\bar{G}_k)_{i,j} = e^{j\theta} (G_k)_{i,j}$ ; thus, Lemma 1 holds and  $A$  and  $B$  have the same eigenvalues. Hence, from (21) we get

$$\lambda_j(G_k^H G_k) + \lambda_{k-j+1}(G_k^H G_k) = \frac{n}{k} = 2. \quad (22)$$

This further implies that for odd values of  $k$  the middle eigenvalue of  $G_k^H G_k$  is 1.

We close this section with an example illustrating some of the above properties. Consider an  $(n, k)$  DFT frame and the the following two cases. First, the rows of  $G_k$  are evenly spaced rows of  $G$  (i.e., either odd rows or even rows). This is the “best” submatrix in the sense that it minimizes the MSE. For such a submatrix, all eigenvalues are known to be equal, as it is a DFT matrix. For example, for  $n = 10, k = 5$ , the best square submatrix results in  $\lambda = 1$  with multiplicity of 5. The other extreme case, which maximizes the MSE, happens when the rows of  $G_k$  are circularly consecutive rows of  $G$ . Again, for the above example,  $\lambda = \{0.0011, 0.1056, 1, 1.8944, 1.9989\}$ . With these examples in mind, we will explore the best and worst frames in Section VII. We shall now discuss signal reconstruction for systematic frames.

## VI. PERFORMANCE ANALYSIS

In this section, we analyze the performance of quantized systematic DFT codes using the quantization model proposed in [10], which assumes that noise components are uncorrelated and each noise component

$q_i$  has mean 0 and variance  $\sigma_q^2$ , i.e., for any  $i, j$ ,

$$\mathbb{E}\{q_i\} = 0, \quad \mathbb{E}\{q_i q_j\} = \sigma_q^2 \delta_{ij}. \quad (23)$$

For one thing,  $q$  can be uniformly distributed on  $[-\Delta/2, \Delta/2]$ , where  $\sigma_q^2 = \Delta^2/12$ . We assume the quantizer range covers the dynamic range of all codewords encoded using the systematic DFT code in (15).

Let  $\mathbf{x}$  be the signal (message) to be transmitted. The corresponding codeword is generated by

$$\mathbf{y} = G_{\text{sys}} \mathbf{x}. \quad (24)$$

This is then quantized to  $\hat{\mathbf{y}}$  and transmitted. Assuming the quantization model in (23), transmitted codeword can be modeled by

$$\hat{\mathbf{y}} = G_{\text{sys}} \mathbf{x} + \mathbf{q}, \quad (25)$$

where  $\mathbf{q}$  represents quantization error. This also models the received codeword provided that there is no error or erasure in channel. Now, suppose we want to estimate  $\mathbf{x}$  from (25). This can be done through the use of linear or nonlinear operations.

#### A. Linear Reconstruction

We first consider *linear reconstruction* of  $\mathbf{x}$  from  $\hat{\mathbf{y}}$  using the pseudoinverse [10] of  $G_{\text{sys}}$ , which is defined by

$$G_{\text{sys}}^\dagger = (G_{\text{sys}}^H G_{\text{sys}})^{-1} G_{\text{sys}}^H = \frac{k}{n} G_k G^H. \quad (26)$$

The linear reconstruction is hence given by

$$\hat{\mathbf{x}} = \frac{k}{n} G_k G^H \hat{\mathbf{y}} = \mathbf{x} + \frac{k}{n} G_k G^H \mathbf{q}, \quad (27)$$

where  $\mathbf{q}$  represents quantization error.

Let us now evaluate the reconstruction error. The mean-squared reconstruction error, due to the quantization noise, using a systematic frame can be written as

$$\begin{aligned} \text{MSE}_q &= \frac{1}{k} \mathbb{E}\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} = \frac{1}{k} \mathbb{E}\{\|G_{\text{sys}}^\dagger \mathbf{q}\|^2\} \\ &= \frac{1}{k} \mathbb{E}\{\mathbf{q}^H G_{\text{sys}}^{\dagger H} G_{\text{sys}}^\dagger \mathbf{q}\} = \frac{1}{k} \sigma_q^2 \text{tr}(G_{\text{sys}}^{\dagger H} G_{\text{sys}}^\dagger) \\ &= \frac{k}{n^2} \sigma_q^2 \text{tr}(G G_k^H G_k G^H) \\ &= \frac{k}{n^2} \sigma_q^2 \text{tr}(G_k^H G_k G^H G) \\ &= \frac{1}{n} \sigma_q^2 \text{tr}(G_k^H G_k) = \frac{k}{n} \sigma_q^2, \end{aligned} \quad (28)$$

where the last step follows because of Lemma 2. This shows that DFT codes reduce quantization error. The MSE is inversely proportional to the redundancy of the frame, which is a well-known result [12], [10].

The above analysis indicates that the MSE is the same for all systematic DFT frames of the same size, no matter they are tight or not. This is however assuming that the effective range codewords generated by different  $G_{\text{sys}}$  is equal, which implies the same  $\sigma_q^2$  for a given number of quantization levels. However, from (17) it is known that, for a fixed number of quantization levels,  $\sigma_q^2$  depends on the variance of transmitted codewords ( $\sigma_y^2$ ) if the quantizer is designed to cover the entire effective range of codewords. Obviously, though,  $\sigma_y^2$  can vary from one systematic frame to another, as shown in (18).

**Theorem 5.** *When encoding with a systematic DFT frame in (15) and decoding with linear reconstruction, for the noise model (23) and given a same number of quantization levels, the MSE is minimum if and only if the systematic frame is tight.*

*Proof:* All systematic DFT frames amount to a same quantization error provided that the effective range of codewords are fully covered, as shown in (28). Nevertheless, for a fixed number of quantization levels more codewords are within the range of quantizer if the systematic frame is tight. This is clear from (18) and (19), recalling that (19) is minimized by the tight frames. Moreover, any frame that minimizes (19) is required to be tight. This will be proved in Section VII-A. ■

The problem we are considering in Theorem 5 is somewhat the dual of that in [12, Theorem 3.1]. Note that in [12, Theorem 3.1] “uniform” frames are used for encoding which implies the same variance for all samples of codewords whereas the reconstruction error is proportional to  $\sum_{i=1}^k \lambda_i$ . On the other hand, the frames ( $G_{\text{sys}}$ ) used in Theorem 5 are not uniform in general; this result in a codeword variance proportional to  $\sum_{i=1}^k \lambda_i$  while having a fixed, minimum reconstruction error.

### B. Consistent Reconstruction

Linear reconstruction is not always the best one can estimate  $\mathbf{x}$  from  $\hat{\mathbf{y}}$ . Although *linear reconstruction* is more tractable, *consistent reconstruction* is known to give significant improvement over linear reconstruction in overcomplete expansions [27]–[29]. Asymptotically, the MSE is  $O(r^{-2})$  for consistent reconstruction, where  $r = n/k$  is the frame redundancy [28]. As it can be seen from (28), for linear reconstruction this is  $O(r^{-1})$ . The improvement, in consistent reconstruction, is due to using deterministic properties of quantization rather than considering quantization as an independent noise as in (23).

Although the MSE in consistent reconstruction is approximated by  $cr^{-2}$ , where the constant  $c$  depends

on the source and quantization, this is verified only if the oversampling ratio  $r$  is very high [29]. In some practical applications of frames, e.g., channel coding, this ratio cannot be high, though. Particularly, in the context of our interest, i.e., DSC,  $r$  is limited to two [15]. Besides, consistent reconstruction methods does not provide a guidance on how to design the frame, as they do not point out how to compute the constant  $c$ . More importantly, (28) proves to be predictive of the performance of consistent reconstruction [10]; therefore, it can be convincingly used as a design criterion regardless of the reconstruction method.

### C. Reconstruction with Error and Erasure

In the context of channel coding, DFT codes are primarily used to provide robustness against channel impairments which can be errors or erasures. Likewise, in DSC these codes play the role of channel codes to combat the errors due to the virtual correlation channel [15]. Thus, it makes sense to evaluate the performance of these codes in the presence of error. To this end, let  $\hat{\mathbf{y}} = G\mathbf{x} + \boldsymbol{\eta}$  where  $\boldsymbol{\eta} = \mathbf{q} + \mathbf{e}$ . Assuming that the quantization and channel errors are independent, we will have

$$\begin{aligned}\mathbb{E}\{\boldsymbol{\eta}^T \boldsymbol{\eta}\} &= \mathbb{E}\{\mathbf{q}^T \mathbf{q} + \mathbf{q}^T \mathbf{e} + \mathbf{e}^T \mathbf{q} + \mathbf{e}^T \mathbf{e}\} \\ &= n\sigma_q^2 + \nu\sigma_e^2,\end{aligned}\tag{29}$$

where  $\nu$  is the average number of errors in each codeword and  $\mathbb{E}\{\mathbf{e}^T \mathbf{e}\} \triangleq \nu\sigma_e^2$ . Note that  $\mathbb{E}\{\mathbf{e}^T \mathbf{q}\} = \mathbb{E}\{\mathbf{q}^T \mathbf{e}\} = 0$ , because  $\mathbf{q}$  and  $\mathbf{e}$  are independent and  $\mathbf{q}$  has mean equal to zero. Finally, following a similar analysis as in (28), we obtain

$$\text{MSE}_{\mathbf{q}+\mathbf{e}} = \frac{k}{n}\sigma_{\boldsymbol{\eta}}^2 = \frac{k}{n}\left(\sigma_q^2 + \frac{\nu}{n}\sigma_e^2\right).\tag{30}$$

From (30) it is clear that reconstruction error has two distinct parts, one due to quantization error and one due to channel error. It also proves that DFT codes decrease both channel and quantization errors by a factor of frame redundancy  $r = n/k$ . The above results is for the case when no error correction is done. It is worth noting that, even without correcting errors, the MSE can be smaller than quantization error.

As another extreme case, let us consider the case when error localization is perfect, i.e., errors are in the *erasure* form. Then, we remove the corrupted samples and do reconstruction using the error-free samples. This approach does not require error correction in order to reconstruct the message; however, it is shown to be equal to the coding theoretic approach [11]. Let  $\hat{\mathbf{y}}_R$  and  $\boldsymbol{\eta}_R$  denote remaining rows of  $\hat{\mathbf{y}}$  and  $\boldsymbol{\eta}$ , respectively. Obviously,  $\boldsymbol{\eta}_R$  includes only quantization error, hence we represent  $\boldsymbol{\eta}_R$  with  $\mathbf{q}_R$ .

Also, let  $F$  denote the rows of  $G_{\text{sys}}$  corresponding to  $\mathbf{q}_R$ . Then, we can write

$$\hat{\mathbf{y}}_R = F\mathbf{x} + \mathbf{q}_R, \quad (31)$$

$$\hat{\mathbf{x}} = F^\dagger \hat{\mathbf{y}}_R, \quad (32)$$

where  $F^\dagger = (F^H F)^{-1} F^H$ . Thus, similar to (28) we will have

$$\begin{aligned} \text{MSE}_{q+\rho} &= \frac{1}{k} \mathbb{E}\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} = \frac{1}{k} \mathbb{E}\{\|F^\dagger \mathbf{q}_R\|^2\} \\ &= \frac{1}{k} \sigma_q^2 \text{tr}(F^\dagger H F^\dagger) \\ &= \frac{1}{k} \sigma_q^2 \text{tr}(F^H F)^{-1} \\ &= \frac{1}{k} \sigma_q^2 \sum_{i=1}^k \frac{1}{\mu_i}, \end{aligned} \quad (33)$$

where subscript  $\rho$  denotes erasure and  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_k > 0$  represent the eigenvalues of  $F^H F$ . We assume at least  $k$  samples are intact which implies  $\mu_k > 0$ .

One nice property of systematic frames is that reconstruction error cannot be more than quantization error as long as systematic samples are intact. This holds even if consecutive samples are erased. We know that consecutive erasures can increase the MSE very fast (e.g., see [11, Table I]). This can be understood from (33) since  $F$  contains  $I_k$  as a subframe and in the worst case we can use this subframe for reconstruction which leads to  $\text{MSE}_{q+\rho} = \sigma_q^2$ . Adding any other row (sample) will decrease the MSE. To show this, let  $F^H = [I_k \mid E^H]$ . Then,  $F^H F = I_k + E^H E$  and, from (8), for  $i = j$ , we get  $\mu_i \geq 1 + \xi_k$  for  $i = 1, \dots, k$ , where  $\xi_k$  is the smallest eigenvalue of  $E^H E$ . Clearly,  $\xi_k \geq 0$  since  $E^H E$  is a positive semidefinite matrix. Further, at least  $\mu_1 > 0$  since otherwise  $E$  must be zero. Hence,  $\sum_{i=1}^k \frac{1}{\mu_i}$  decreases by adding new rows.

Finally, with consistent reconstruction, we can further decrease the MSE. To do so, we check if reconstructed values  $\hat{x}_i$  for systematic samples in (32) are consistent with their values before reconstruction or not, i.e., for any systematic sample, we must have  $Q(\hat{x}_i) = Q(\hat{y}_{Ri})$ . Otherwise, we replace  $\hat{x}_i$  with

$$\hat{\hat{x}}_i = Q(\hat{y}_{Ri}) - \text{sign}(Q(\hat{y}_{Ri}) - \hat{x}_i) \frac{\Delta}{2}. \quad (34)$$

## VII. CLASSIFICATION OF SYSTEMATIC FRAMES

### A. The Best and Worst Systematic Frames

As we discussed in Section V, the optimal  $G_{\text{sys}}$  is achieved from the optimization problem (19). Similarly, to find the worst  $G_{\text{sys}}$ , we can *maximize* (19) instead of minimizing it. The optimal eigenvalues



are known to be  $\lambda_i = 1, 1 \leq i \leq k$ . But, how can we find the corresponding  $G_{\text{sys}}$ , or  $G_k$  equivalently? More importantly, if a  $G_k$  with  $\lambda_i = 1$  does not exist, is there any suggestion for the best matrix?

We approach this problem by studying another optimization problem. To this end, we first prove the following theorem for the eigenvalues of a matrix.

**Theorem 6.** *Let  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$  be the eigenvalues of a nonsingular  $k \times k$  matrix  $A$ , then we have*

$$\left(\sum_{i=1}^k \frac{1}{\lambda_i}\right) \cdot \left(\prod_{i=1}^k \lambda_i\right) = c, \quad (35)$$

where the constant  $c$  is a function of  $\text{tr}(A), \dots, \text{tr}(A^{k-1})$ .

*Proof:* See Section IX-B. ■

Now, in view of Theorem (6), we can see that

$$\underset{\lambda_i}{\operatorname{argmin}} \sum_{i=1}^k \frac{1}{\lambda_i} = \underset{\lambda_i}{\operatorname{argmax}} \prod_{i=1}^k \lambda_i. \quad (36)$$

As a result, the optimal arguments of the optimization problem in (19) are equal to those of

$$\begin{aligned} & \underset{\lambda_i}{\operatorname{maximize}} \quad \prod_{i=1}^k \lambda_i \\ & \text{s.t.} \quad \sum_{i=1}^k \lambda_i = k, \quad \lambda_i > 0, \end{aligned} \quad (37)$$

in which  $\{\lambda_i\}_{i=1}^k$  are the eigenvalues of  $G_k G_k^H$  (or  $V_k^H V_k$ ). By using the Lagrangian method, one can check that (37) has the maximum of 1 and infimum of 0. Then, considering that

$$\prod_{i=1}^k \lambda_i = \det(V_k^H V_k) = \det(G_k G_k^H), \quad (38)$$

we conclude that the “best” submatrix is the one with the largest determinant (possibly 1) and the “worst” submatrix is the one with smallest determinant.

Next, we evaluate the determinant of  $V_k^H V_k$  so as to find the matrices corresponding to the extreme cases. To this end, we first evaluate the determinate of  $WW^H$  where  $W$  is the Vandermonde matrix with unit complex entries as defined in (2). From (3) we can write

$$\begin{aligned} \det(WW^H) &= \frac{1}{n^n} \prod_{1 \leq p < q \leq n} |e^{i\theta_p} - e^{i\theta_q}|^2 \\ &= \frac{1}{n^n} \prod_{1 \leq p < q \leq n} 4 \sin^2 \frac{\pi}{n} (q - p) \\ &= \frac{2^{n(n-1)}}{n^n} \prod_{r=1}^{n-1} \left( \sin^2 \frac{\pi}{n} r \right)^{n-r}, \end{aligned} \quad (39)$$

in which  $\theta_x = \frac{2\pi}{n}(x-1)$ ,  $r = q-p$ , and  $n(n-1)/2$  is the total number of terms that satisfy  $1 \leq p < q \leq n$ . But, we see that  $W$  is a DFT matrix, and thus, its determinant must be 1. Therefore, we have

$$\prod_{r=1}^{n-1} \left( \sin^2 \frac{\pi}{n} r \right)^{n-r} = \frac{n^n}{2^{n(n-1)}}. \quad (40)$$

The above analysis helps us evaluate the determinant of  $V_k$  or  $G_k$ , defined in (16). Let  $\mathcal{I}_{r_k} = \{i_{r_1}, i_{r_2}, \dots, i_{r_k}\}$  be those rows of  $G$  used to build  $G_k$ . Also, without loss of generality, assume  $i_{r_1} < i_{r_2} < \dots < i_{r_k}$ . Clearly,  $i_{r_1} \geq 1, i_{r_k} \leq n$ , and we obtain

$$\begin{aligned} \det(V_k V_k^H) &= \frac{1}{k^k} \prod_{\substack{1 \leq p < q \leq n \\ p, q \in \mathcal{I}_{r_k}}} |e^{i\theta_p} - e^{i\theta_q}|^2 \\ &= \frac{1}{k^k} \prod_{\substack{1 \leq p < q \leq n \\ p, q \in \mathcal{I}_{r_k}}} 4 \sin^2 \frac{\pi}{n} (q - p). \end{aligned} \quad (41)$$

Then, since  $\sin \frac{\pi}{n} u = \sin \frac{\pi}{n} (n - u)$ , one can see that this determinant depends on the circular distance between rows in  $\mathcal{I}_{r_k}$ . For a matrix with  $n$  rows, we define the circular distance between rows  $p$  and  $q$  as  $\min \{|q - p|, n - |q - p|\}$ . In this sense, for example, the distance between rows 1 and  $n$  is one, i.e., they are circularly successive. Now, it is reasonable to believe that (41) is minimized when the selected rows are (circularly) successive. Note that  $\sin u$  is strictly increasing for  $u \in [0, \pi/2]$ , and the circular distance cannot be greater than  $n/2$  in this problem.

In such circumstances where all rows in  $\mathcal{I}_{r_k}$  are (circularly) successive, (41) is minimal and reduces to

$$\det(V_k V_k^H) = \frac{2^{k(k-1)}}{k^k} \prod_{r=1}^{k-1} \left( \sin^2 \frac{\pi}{n} r \right)^{k-r}. \quad (42)$$

The other extreme case comes up when  $n = Mk$  ( $M$  is a positive integer) provided that  $G_k$  consists of every  $M$ th row of  $G$ . In such a case, (41) simplifies to 1, because

$$\begin{aligned} \det(V_k V_k^H) &= \frac{2^{k(k-1)}}{k^k} \prod_{r=1}^{k-1} \left( \sin^2 \frac{\pi}{n} Mr \right)^{k-r} \\ &= \frac{2^{k(k-1)}}{k^k} \prod_{r=1}^{k-1} \left( \sin^2 \frac{\pi}{k} r \right)^{k-r} \\ &= 1, \end{aligned} \quad (43)$$

where the last step is due to (40). Recall that this gives the best  $V_k$  (and equivalently  $G_k$ ), in light of (37). For such a  $G_k$ , it is easy to see that  $G_{\text{sys}}$  stands for a “tight” systematic frame and minimizes the MSE

for a given number of quantization levels. Effectively, such a frame is performing *integer oversampling*. There are  $M$  such frames; they all have the same spectrum, though.

Recall that, from the optimization problem in (36)–(38) and Theorem 3,  $\det(V_k V_k^H) < 1$  for  $n \neq Mk$ . For such an  $(n, k)$  frame, systematic rows cannot be equally spaced in the corresponding systematic frame; instead, we may explore a systematic frame in which the circular distance between successive systematic samples is as evenly as possible. Then, the circular distance between each successive systematic rows is either  $\lfloor n/k \rfloor$  or  $\lceil n/k \rceil$ . More precisely, if  $l$  and  $m$ , respectively, represent the number of systematic rows with circular distance equal to  $\lceil n/k \rceil$  and  $\lfloor n/k \rfloor$ , they must satisfy

$$\begin{cases} l + m = k, \\ l \lceil \frac{n}{k} \rceil + m \lfloor \frac{n}{k} \rfloor = n. \end{cases} \quad (44)$$

In the following theorem, we prove that the best performance is achieved when the systematic rows are as equally spaced as possible, i.e., when (44) is satisfied.

**Theorem 7.** *When encoding with an  $(n, k)$  systematic DFT frame in (15) and decoding with linear reconstruction, for the noise model (23) and given a same number of quantization levels, the MSE is minimum when there are  $l = n - \lfloor n/k \rfloor k$  systematic rows with successive circular distance  $\lceil n/k \rceil$  while the remaining  $m = k - l$  systematic rows have a successive circular distance equal to  $\lfloor n/k \rfloor$ .*

*Proof:* See Appendix IX-C. ■

Effectively, the above theorem is generalizing Theorem 5. Note that when  $n = Mk$ ,  $\lfloor n/k \rfloor = \lceil n/k \rceil = M$  and there exist  $k$  systematic rows with equal distance; in this case, Theorem 7 reduces to Theorem 5 and the corresponding systematic frame is tight. The optimality of this case was proved in (43). When  $n \neq Mk$ , we cannot have a systematic frame with equally spaced systematic rows; however, the best performance is still achieved when the circular distance between systematic rows is as evenly as possible, as explained above. Note that in either case  $d_{min}$ , the minimum distance between systematic rows, is  $\lfloor n/k \rfloor$ . This is a necessary condition for an optimal systematic frame, as shown in the proof of Theorem 7.

### B. Numerical Examples

Numerical calculations confirm that “evenly” spaced data samples gives rise to systematic frames with the best performance. When a systematic code is performing integer oversampling, we end up with tight systematic frames. The first and last codes in Table I are examples of this case. When  $n \neq Mk$ , data samples cannot be equally spaced; however, as it can be seen from the second code in Table I, still the

TABLE I  
EIGENVALUES STRUCTURE FOR TWO SYSTEMATIC DFT FRAMES WITH DIFFERENT CODEWORD PATTERNS. A “ $\times$ ” AND “ $-$ ”  
RESPECTIVELY REPRESENT DATA (SYSTEMATIC) AND PARITY SAMPLES.

Code	Codeword pattern	$\lambda_{\min}$	$\lambda_{\max}$	$\sum_{i=1}^k 1/\lambda_i$	$\prod_{i=1}^k \lambda_i$
(6, 3)	$\times \times \times - - -$	0.0572	1.9428	19	0.1111
	$\times \times - \times - -$	0.2546	1.7454	5.5	0.4444
	$\times \times - - \times -$	0.2546	1.7454	5.5	0.4444
	$\times - \times - \times -$	1	1	3	1
(7, 5)	$\times \times \times \times \times - -$	0.0396	1.4	28.70	0.0827
	$\times \times \times \times - \times -$	0.1506	1.4	10.32	0.2684
	$\times \times - \times \times - \times$	0.3110	1.4	7.40	0.4173
	$\times - \times \times \times - \times$	0.3110	1.4	7.40	0.4173
(10, 5)	$\times \times \times \times \times - - - -$	0.0011	1.9989	908.21	$4.46 \times 10^{-4}$
	$\times \times \times \times - \times - - -$	0.0041	1.9959	249.94	0.0047
	$\times \times \times \times - - \times - -$	0.0110	1.9890	96.09	0.0122
	$\times \times \times - - \times \times - -$	0.0496	1.9504	25.64	0.0489
	$\times \times - \times \times - \times - -$	0.0512	1.9488	23.41	0.0838
	$\times \times - \times \times - - \times -$	0.1056	1.8944	13.79	0.1436
	$\times - \times - \times - \times - - \times$	0.2377	1.7623	7.77	0.4189
	$\times - \times - \times - \times - \times -$	1	1	5	1

best performance is achieved when they are as equally spaced as possible. In this table, “ $\times$ ’s” and “ $-$ ’s” represent data and parity samples, respectively. Moreover, we observe that, circularly shifted codeword patterns behave the same (e.g., in the (7, 5) code, frames with pattern  $\times - \times \times \times - \times$  and  $\times \times - \times \times - \times$  have the same performance). Also, reversal of a codeword pattern yields a codeword with the same performance (e.g.,  $\times \times - \times - -$  is shifted version of reversed  $\times \times - - \times -$  in the (6, 3) code). These properties hold in general, as stated below.

**Property 1.** *Circular shift of  $\mathcal{I}_{r_k}$ , the systematic rows of a systematic frame with analysis frame  $G_{\text{sys}}$ , does not change the spectrum of  $G_{\text{sys}}^H G_{\text{sys}}$ .*

**Property 2.** *Reversal of  $\mathcal{I}_{r_k}$  yields a systematic frame with the same spectral properties.*

*Proof:* From (15) we obtain

$$\lambda_i(G_{\text{sys}}^H G_{\text{sys}}) = \frac{n/k}{\lambda_i(G_k G_k^H)}, \quad (45)$$

for  $i = 1, \dots, k$ . But  $G_k G_k^H$  is invariant to the circular shift of rows of  $G$  that make  $\mathcal{I}_{r_k}$ , as long as all rows are shifted the same amount in the same direction. This can be seen from the proof of Lemma 2 in (14) by defining  $r' = r + c$  where  $r'$  represent the shifted rows by a constant  $c$  and  $r \in \mathcal{I}_{r_k}$ . This proves Property 1. Likewise, let  $r'' = n + 1 - r$  be the reversed row indices. Again, from (14), it is clear that Property 2 holds. ■

These properties together show that the frame operators of systematic frames ( $G_{\text{sys}}^H G_{\text{sys}}$ ), in which the “relative” circular distance among the systematic rows are the same, inherit the same spectrum and thus show the same performance.

### C. Number of Systematic Frames

The number of systematic frames is obviously finite but their performance depends on the position of systematic rows, or equivalently, the position of data (or parity) samples in the associated codewords, and can be the same for different systematic frames. In what follows, we derive an upper and lower bound on the number of systematic frames with different spectrum. In other words, we categorize these frames based on their performance. To this end, we observe that the problem of finding  $k \times k$  submatrices of an  $n \times k$  matrix can be viewed as finding different  $k$ -subsets of a set with  $n$  elements. This is given by the binomial coefficient  $\binom{n}{k}$  and is also equivalent to the number of systematic frames. As stated earlier in Property 1, circular shift of a codeword pattern does not change its spectrum, and so its performance. We define a *coset* as square submatrices that result in a same performance. Each coset has at least  $n$  elements ( $k$ -subsets), as shown in Table II. To find these elements, it suffices to circularly shift a subset  $n$  times. Equivalently, for a given  $k$ -subset, we simply add up 1 to each element of a subset. Note that, the subsets elements are  $k$  row indices of  $G_{n \times k}$  and thus cannot be greater than  $n$ . Therefore, once a shifted index  $x$  becomes greater than  $n$ , we replace it with  $\langle\langle x \rangle\rangle_n$  where  $\langle\langle x \rangle\rangle_n \triangleq x - dn$  if  $dn + 1 \leq x \leq dn + n, d \in \mathbb{Z}$ . Obviously, each coset has at least  $n$  subsets since  $n - 1$  circular shifts of a given subset are distinct; all these subsets have the same relative distance, though. This can be seen in Table II. Thus, it is clear that the number of cosets is the bounded by

$$n_c \leq u = \frac{1}{n} \binom{n}{k}. \quad (46)$$

TABLE II  
DIFFERENT COSETS OF  $(7, 3)$  DFT FRAME AND THEIR CORRESPONDING RELATIVE DISTANCES AND SPECTRUMS. THE  
COSET LEADERS ARE IN BOLDFACE.

	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>
Leader	<b>1 2 3</b>	<b>1 2 4</b>	<b>1 2 5</b>	<b>1 3 5</b>	<b>1 3 4</b>
	2 3 4	2 3 5	2 3 6	2 4 6	2 4 5
	3 4 5	3 4 6	3 4 7	3 5 7	3 5 6
	4 5 6	4 5 7	1 4 5	1 4 6	4 6 7
	5 6 7	1 5 6	2 5 6	2 5 7	1 5 7
	1 6 7	2 6 7	3 6 7	1 3 6	1 2 6
	1 2 7	1 3 7	1 4 7	2 4 7	2 3 7
Distance	1 1 2	1 2 3	1 3 3	2 2 3	1 3 2
Weight	4	6	7	7	6
$\lambda_1$	2.1558	1.7539	1.9066	1.2673	1.7539
$\lambda_2$	0.8150	1.1133	0.8424	1.1601	1.1133
$\lambda_3$	0.0292	0.1328	0.2510	0.5726	0.1328

Let  $\mathcal{I}_{r_k}^r$  denote the reversal of  $\mathcal{I}_{r_k} = \{i_{r_1}, i_{r_2}, \dots, i_{r_k}\}$ . We define

$$\mathcal{I}_{r_k}^r = \langle\langle n + 1 - \mathcal{I}_{r_k} \rangle\rangle_n. \quad (47)$$

This operation is performed on every element of  $\mathcal{I}_{r_k}$ . One can see that reversal of a subset does not change its distance and spectrum, owing to Property 2. This can reduce the number of cosets. For example, in Table II, the reversal of  $\{1, 2, 4\}$ , which is the *coset leader* in  $C_2$ , is  $\{7, 6, 4\}$  which belongs to  $C_5$ . This indicates  $C_2$  and  $C_5$  are essentially one coset. The bound in (46) is tight if and only if there are  $u$  self-reversal cosets. Trivial examples of such a code are achieved when  $k = n - 1$  or  $k = 1$ . A self-reversal coset is a coset that the reversal of its elements belong to itself, e.g.,  $C_1$ ,  $C_3$ , and  $C_4$  in Table II.

On the other hand,  $n_c \geq u/2$  is a lower bound because there cannot be more than one reversal for a given coset. It can be further seen that the coset with smallest weight ( $C_1$ ) is always self-reverse, i.e., the reversal of each element of  $C_1$  is its own element for any  $(n, k)$  frame. This implies that the lower bound is not achievable. Therefore,

$$\frac{1}{2n} \binom{n}{k} < n_c \leq \frac{1}{n} \binom{n}{k}. \quad (48)$$

## VIII. CONCLUSIONS

We have introduced, proposed the construction method, and analyzed the performance of systematic DFT frames in this paper. Several systematic DFT frames can be made out of the generator matrix of a BCH-DFT code. The performance of these frames, however, depends on the relative position of systematic samples or, equivalently, parity samples in the codeword. We proved that evenly spaced systematic (parity) samples result in the minimum mean-squared reconstruction error, whereas the worst performance is expected when systematic samples are circularly consecutive. Further, we found the conditions for which a systematic DFT frame can be tight, too. Tight systematic DFT frame can be realized only if the frame is performing integer oversampling and systematic samples are circularly equally spaced. Finally, for each DFT frame, we classified systematic DFT frame based on their performance.

It would be interesting to extend this work to oversampled DFT filter banks, an infinite-dimension of DFT frames. Oversampled filter banks can be used as error correcting codes [30]. In a relevant work, systematic wavelet subcodes, a form of real-number convolutional codes, has been proposed for data protection [31].

## ACKNOWLEDGEMENT

The authors wish to thank an anonymous reviewer of the conference version of this work for several valuable comments. We also thank Sina Hamidi Ghalehjegh and Mohsen Akbari for fruitful discussions and their comments.

## IX. APPENDIX

### A. Proof of Theorem 3

*Proof:* Let  $n = Mk + l$ ,  $0 < l < k$ , then  $G$  can be partitioned as  $G = [G_k^H \mid G_k^{1H} \mid \dots \mid G_k^{(M-1)H} \mid G_{k \times l}^{MH}]^H$ . In general,  $G_k, G_k^1, \dots, G_k^{M-1}$  and  $G_{k \times l}^M$  include arbitrary rows of  $G$ , hence they have different spectrums, i.e., different sets of eigenvalues. Suppose, for the purpose of contradiction, that  $\lambda_k(G_k^H G_k) = 1$ ; this can occur only if  $G_k$  consist of the rows of  $G$  such that the distance between each two successive rows is at least  $M$ .<sup>5</sup> Such an arrangement guarantees the existence of  $G_k^1, \dots, G_k^{M-1}$  so that  $G_k^{mH} G_k^m$ , for any  $1 \leq m \leq M-1$ , has the same spectrum as  $G_k^H G_k$ . To find the row indices corresponding to  $G_k^m$ , we can simply add  $m$  to each row index of  $G_k$ . Then, to show these matrices have the same spectrum, we use Lemma 1. Given a  $G_k$ , one can verify that  $(G_k^m)_{i,j} = e^{j\frac{2\pi m}{n}}(G_k)_{i,j}$  and thus  $(G_k^m)^H_{i,j} = e^{-j\frac{2\pi m}{n}}(G_k)^H_{i,j}$ .

<sup>5</sup> $\lambda_k(G_k^H G_k) = 1$  is the optimal solution for (19) and necessitate  $d_{min} = M$ , as discussed in Theorem 7.

Therefore,  $G_k^{mH} G_k^m$  and  $G_k^H G_k$  have the same spectrum for any  $1 \leq m \leq M-1$ . Next, we see that  $G^H G = A + B$  in which  $A = G_k^H G_k + \dots + G_k^{(M-1)H} G_k^{M-1}$  and  $B = G_{k \times l}^{MH} G_{k \times l}^M$ . Then, in consideration of the above discussion,  $\lambda_i(A) = M\lambda_i(G_k^H G_k)$  for any  $1 \leq i \leq k$ . Hence, from (8), for  $i = 1, j = k$ , we will have

$$\begin{aligned} \lambda_k(A) + \lambda_1(B) &\leq \lambda_1(A + B) \\ \Leftrightarrow M\lambda_k(G_k^H G_k) &\leq \frac{n}{k} - \lambda_1(B) \\ \Leftrightarrow \lambda_k(G_k^H G_k) &\leq \frac{\frac{n}{k} - 1}{M} = \frac{\frac{n}{k} - 1}{\lfloor \frac{n}{k} \rfloor} < 1, \end{aligned} \quad (49)$$

where the last line follows using  $\lambda_1(B) \geq 1$  from Theorem 1. But this is contradicting our assumption  $\lambda_k(G_k^H G_k) = 1$ , and thus completes the proof that, for  $n \neq Mk$ , the largest possible  $\lambda_k(G_k^H G_k)$  is strictly less than 1, for any  $G_k$ .<sup>6</sup>

The proof of the other bound ( $\lambda_1(G_k^H G_k) > 1$ ) is then immediate because  $\sum_{i=1}^k \lambda_i(G_k^H G_k) = \sum_{i=1}^k a_{ii} = k$ . ■

### B. Proof of Theorem 6

*Proof:* Let  $A$  be a  $k \times k$  matrix. The eigenvalues of  $A$  are the roots of the *characteristic polynomial* defined as

$$P(\lambda) = \prod_{i=1}^k (\lambda - \lambda_i). \quad (50)$$

It can be shown that

$$\frac{P'(\lambda)}{P(\lambda)} = \sum_{i=1}^k \frac{1}{\lambda - \lambda_i}. \quad (51)$$

Next, by setting  $\lambda = 0$  we get

$$\left( \sum_{i=1}^k \frac{1}{\lambda_i} \right) \cdot \left( \prod_{i=1}^k \lambda_i \right) = (-1)^{k+1} P'(0) \triangleq c. \quad (52)$$

Clearly  $P'(0)$  is a constant for a given matrix  $A$ ; it is a function of eigenvalues of  $A$ . To find this constant, we use Newton's identities to re-express the characteristic polynomial in (50) as

$$P(\lambda) = p_0 \lambda^k + p_1 \lambda^{k-1} + \dots + p_{k-1} \lambda + p_k, \quad (53)$$

<sup>6</sup>Note that when  $n = Mk$ ,  $B$  is an empty matrix and we must put  $\lambda_1(B) = 0$  in (49) which result in  $\lambda_k(G_k^H G_k) \leq 1$  and does not guarantee a bound strictly less than 1.



in which  $p_0 = 1$  and, for  $j = 1, \dots, k$ ,

$$p_j = -\frac{1}{j} \sum_{l=0}^{j-1} p_l s_{j-l}, \quad (54)$$

$$s_j = \sum_{i=1}^k \lambda_i^j = \text{tr}(A^j). \quad (55)$$

Thus, the constant  $c$  in (52) is given by  $c = (-1)^{k+1} p_{k-1}$  which is a function of  $\text{tr}(A), \dots, \text{tr}(A^{k-1})$ . ■

### C. Proof of Theorem 7

*Proof:* Consider an  $(n, k)$  DFT frame, let  $M = \lfloor n/k \rfloor$ , and assume that all rows in  $\mathcal{I}_{r_k}$ , except the first and last rows, are equally spaced with distance  $M$  (without loss of generality, we assume  $i_{r_1} = 1$ , then  $i_{r_j} = (j-1)M + 1$ ,  $j \leq k$ ). Hence  $d_{\min} = M$ , where the minimum distance  $d_{\min}$  is defined as the smallest circular distance among the selected rows. In such a setting, from (41) and similar to (43), we can write

$$\det(V_k V_k^H) = \frac{2^{k(k-1)}}{k^k} \prod_{r=1}^{k-1} \left( \sin^2 \frac{\pi}{n} M r \right)^{k-r}. \quad (56)$$

We prove that, in view of (37), the systematic frame corresponding to the above arrangement has better performance than any other arrangement in which the minimum distance among the systematic rows is less than  $M$ . To this end, we first assume that all selected rows in  $\mathcal{I}_{r_k}$  remain the same except one row which is shifted one unit in a way that  $d_{\min}$  decreases. For example, without loss of generality, consider  $\mathcal{I}'_{r_k}$  for which  $i'_{r_1} = 2$ ,  $i'_{r_j} = i_{r_j}$ ,  $1 < j \leq k$ ; hence  $d_{\min} = M - 1$ . Then, from (41), we obtain

$$\frac{\det(V_k V_k^H)|_{\mathcal{I}'_{r_k}}}{\det(V_k V_k^H)|_{\mathcal{I}_{r_k}}} = \frac{\prod_{r=1}^{k-1} \sin^2 \frac{\pi}{n} (M r - 1)}{\prod_{r=1}^{k-1} \sin^2 \frac{\pi}{n} M r} < 1. \quad (57)$$

To prove the inequality, equivalently, we show that

$$\frac{\sin \frac{(M-1)\pi}{n} \sin \frac{(2M-1)\pi}{n} \dots \sin \frac{((k-1)M-1)\pi}{n}}{\sin \frac{M\pi}{n} \sin \frac{2M\pi}{n} \dots \sin \frac{(k-1)M\pi}{n}} < 1. \quad (58)$$

We break up this inequality into  $\lfloor k/2 \rfloor$  inequalities, each of which strictly less than one. First, consider the first and last terms in the numerator and denominator. We can write

$$\frac{\sin \frac{(M-1)\pi}{n} \sin \frac{((k-1)M-1)\pi}{n}}{\sin \frac{M\pi}{n} \sin \frac{(k-1)M\pi}{n}} = \frac{\cos \frac{(k-2)M\pi}{n} - \cos \frac{(kM-2)\pi}{n}}{\cos \frac{(k-2)M\pi}{n} - \cos \frac{kM\pi}{n}} < 1, \quad (59)$$

where the inequality follows since  $\cos \frac{(kM-2)\pi}{n} > \cos \frac{kM\pi}{n}$ , as  $\frac{kM}{n}\pi \leq \pi$ . Likewise, for the second and penultimate terms we have

$$\frac{\sin \frac{(2M-1)\pi}{n} \sin \frac{((k-2)M-1)\pi}{n}}{\sin \frac{2M\pi}{n} \sin \frac{(k-2)M\pi}{n}} = \frac{\cos \frac{(k-4)M\pi}{n} - \cos \frac{(kM-2)\pi}{n}}{\cos \frac{(k-4)M\pi}{n} - \cos \frac{kM\pi}{n}} < 1. \quad (60)$$

A similar reasoning can be used for other terms that are equally spaced from the two ends.

Clearly, the same argument is valid when  $2 < i'_{r_1} < M$  and the other rows are the same, i.e.,  $i'_{r_j} = i_{r_j}, 1 < j \leq k$  and  $d_{min} = M - i'_{r_1}$ . Moreover, when more than one row index is changed, in a way that two or more selected rows have a distance less than  $M$ , the above argument is valid and we can show that new determinant is even less than the case with one changed index. In fact, in such a case, it is easier to compare the new one with its parent; i.e., to compare the case with two changes with the case with one change. As a result, we can see that any combination of rows with  $d_{min} < M$  performs worse than the case with  $d_{min} = M$ , on account of (37); that is,  $d_{min} = M$  is necessary condition for optimality. In other words, that optimal systematic frame must satisfy  $d_{min} = M$ .

Next, we show that among systematic frames with  $d_{min} = M$  the one that satisfies (44) is the best. That is, the optimal systematic frame has  $l = n - \lfloor n/k \rfloor k$  systematic rows with successive circular distance of  $\lceil n/k \rceil$  and  $m = k - l$  systematic rows with successive circular distance of  $\lfloor n/k \rfloor$ . To prove this, again we compare  $\det(V_k V_k^H)$  in (41) for this case and the other cases with  $d_{min} = M$ . The arguments are very similar to what we used above. Before moving on, we should mention that for  $l \in \{0, 1, k-1\}$  the proof in the first part is sufficient.

Let  $\mathcal{I}_{r_k}^o$  denote the set of rows satisfying the constraints in (44); obviously,  $d_{min} = M$ . We claim that any other selection of systematic rows, for which  $d_{min}$  is  $M$ , results in a smaller  $\det(V_k V_k^H)$ ; that is,  $\det(V_k V_k^H)|_{\mathcal{I}_{r_k}} < \det(V_k V_k^H)|_{\mathcal{I}_{r_k}^o}$ . Let us evaluate the case where only the row index for one of those  $l$  rows varies, provided that  $d_{min} = M$  is kept.<sup>7</sup> We then have

$$\frac{\det(V_k V_k^H)|_{\mathcal{I}_{r_k}}}{\det(V_k V_k^H)|_{\mathcal{I}_{r_k}^o}} = \frac{\prod_{r=1}^{k-1} \sin^2 \frac{\pi}{n} M r}{\prod_{r=1}^{k-1} \sin^2 \frac{\pi}{n} (M r + 1)} < 1. \quad (61)$$

Again it suffice to prove that

$$\frac{\sin \frac{M\pi}{n} \sin \frac{2M\pi}{n} \cdots \sin \frac{(k-1)M\pi}{n}}{\sin \frac{(M+1)\pi}{n} \sin \frac{(2M+1)\pi}{n} \cdots \sin \frac{((k-1)M+1)\pi}{n}} < 1, \quad (62)$$

<sup>7</sup> Note that, with this shift of row, we are looking for an arrangement of a systematic frame that does not satisfy (44); otherwise,  $\det(V_k V_k^H)$  will not vary, as the frame properties has not changed essentially. More specifically, a new, different arrangement will introduce a new distance equal to  $\lceil n/k \rceil + 1$ .

and this can be done by the same divide and conquer approach, used in the first part of this proof. For instance, for the first and last terms in the numerator and denominator we have

$$\frac{\sin \frac{M\pi}{n} \sin \frac{(k-1)M\pi}{n}}{\sin \frac{(M+1)\pi}{n} \sin \frac{((k-1)M+1)\pi}{n}} = \frac{\cos \frac{(k-2)M\pi}{n} - \cos \frac{kM\pi}{n}}{\cos \frac{(k-2)M\pi}{n} - \cos \frac{(kM+2)\pi}{n}} < 1, \quad (63)$$

where the inequality follows for  $\cos \frac{(kM+2)\pi}{n} < \cos \frac{kM\pi}{n}$ . Finally, the other cases, where two or more rows change, can be proved comparing their determinant with their ancestor's with a similar reasoning. This completes the proof that a systematic frame with the most evenly spaced systematic rows, or equivalently data samples in the corresponding codewords, is the best in the minimum MSE sense. ■

## REFERENCES

- [1] M. Vaezi and F. Labeau, "Systematic DFT frames: Principle and eigenvalues structure," in *Proc. International Symposium on Information Theory (ISIT)*, pp. 2436–2440, 2012.
- [2] R. E. Blahut, *Algebraic Codes for Data Transmission*. New York: Cambridge Univ. Press, 2003.
- [3] T. Marshall Jr., "Coding of real-number sequences for error correction: A digital signal processing problem," *IEEE Journal on Selected Areas in Communications*, vol. 2, pp. 381–392, March 1984.
- [4] J. K. Wolf, "Redundancy, the discrete Fourier transform, and impulse noise cancellation," *IEEE Transactions on Communications*, vol. 31, pp. 458–461, March 1983.
- [5] F. Marvasti, M. Hasan, M. Echhart, and S. Talebi, "Efficient algorithms for burst error recovery using FFT and other transform kernels," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1065–1075, April 1999.
- [6] G. Rath and C. Guillemot, "Subspace algorithms for error localization with quantized DFT codes," *IEEE Transactions on Communications*, vol. 52, pp. 2115–2124, Dec. 2004.
- [7] G. Rath and C. Guillemot, "Subspace-based error and erasure correction with DFT codes for wireless channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 3241–3252, Nov. 2004.
- [8] A. Gabay, M. Kieffer, and P. Duhamel, "Joint source-channel coding using real BCH codes for robust image transmission," *IEEE Transactions on Image Processing*, vol. 16, pp. 1568–1583, June 2007.
- [9] G. Takos and C. N. Hadjicostis, "Determination of the number of errors in DFT codes subject to low-level quantization noise," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1043–1054, March 2008.
- [10] V. K. Goyal, J. Kovačević, and J. A. Kelner, "Quantized frame expansions with erasures," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 203–233, 2001.
- [11] G. Rath and C. Guillemot, "Frame-theoretic analysis of DFT codes with erasures," *IEEE Transactions on Signal Processing*, vol. 52, pp. 447–460, Feb. 2004.
- [12] J. Kovačević and A. Chebira, "Life beyond bases: The advent of frames (Part I)," *IEEE Signal Processing Magazine*, vol. 24, pp. 86–104, July 2007.
- [13] P. Casazza and J. Kovačević, "Equal-norm tight frames with erasures," *Advances in Computational Mathematics*, vol. 18, no. 2, pp. 387–430, 2003.

- [14] J. Kovačević and A. Chebira, *An introduction to frames*. Now Publishers, 2008.
- [15] M. Vaezi and F. Labeau, “Distributed lossy source coding using real-number codes,” *to appear in VTC-Fall2012*. [Online]. Available: <http://arxiv.org/abs/1111.0654>.
- [16] P. L. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms, and Applications*. Academic Press, 2009.
- [17] S. K. Mitra and Y. Kuo, *Digital Signal Processing: A Computer-Based Approach*. New York: McGraw-Hill, 2006.
- [18] G. H. Tucci and P. A. Whiting, “Asymptotic results on generalized vandermonde matrices and their extreme eigenvalues,” in *Proc. the 49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1816–1823, 2011.
- [19] G. H. Tucci and P. A. Whiting, “Eigenvalue results for large scale random vandermonde matrices with unit complex entries,” *IEEE Transactions on Information Theory*, vol. 57, pp. 3938–3954, June 2011.
- [20] G. A. F. Seber, *A Matrix Handbook for Statisticians*. New Jersey: John Wiley & Sons, 2008.
- [21] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Now Publishers, 2006.
- [22] J. Bajcsy and P. Mitran, “Coding for the Slepian-Wolf problem with turbo codes,” in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, pp. 1400–1404, 2001.
- [23] J. Garcia-Frias and Y. Zhao, “Compression of correlated binary sources using turbo codes,” *IEEE Communications Letters*, vol. 5, pp. 417–419, Oct. 2001.
- [24] A. Aaron and B. Girod, “Compression with side information using turbo codes,” in *Proc. IEEE Data Compression Conference*, pp. 252–261, 2002.
- [25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 2006.
- [26] T. Chen, E. Serpedin, and D. Rajan, *Mathematical Foundations for Signal Processing, Communications, and Networking*. CRC Press, 2011.
- [27] N. T. Thao and M. Vetterli, “Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 519–531, March 1994.
- [28] N. T. Thao and M. Vetterli, “Reduction of the MSE in  $R$ -times oversampled A/D conversion  $O(1/R)$  to  $O(1/R^2)$ ,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 200–203, Jan. 1994.
- [29] V. K. Goyal, M. Vetterli, and N. T. Thao, “Quantized overcomplete expansions in  $\mathbb{R}^N$ : Analysis, synthesis, and algorithms,” *IEEE Transactions on Information Theory*, vol. 44, pp. 16–31, Jan. 1998.
- [30] F. Labeau, J. C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Macq, “Oversampled filter banks as error correcting codes: Theory and impulse noise correction,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 4619–4630, Dec. 2005.
- [31] G. Redinbo, “Systematic wavelet subcodes for data protection,” *IEEE Transactions on Computers*, vol. 60, pp. 904–909, June 2011.